A Genetic Algorithm for Trip Distribution and Traffic Assignment from Traffic Counts in a Stochastic User Equilibrium

Kiseok Sung (Corresponding Author)
Professor
Industrial and Systems Engineering Department
Kangnung National University
GangNeung, 210-702, South Korea.
E-mail:sung@kangnung.ac.kr
Tel: +82 (33) 640-2371
Fax: +82 (33) 640-2244

To be considered for the publication in IJMS and presentation in KORMS 2006

September 31, 2006

사용자 평형을 이루는 통행분포와 통행배정을 위한 유전알고리즘.

혼잡한 교통네트워크에서 조사된 통행량으로부터 확률적 사용자 평형을 이루는 통행분포와 통행배정을 동시에 구하기 위한 네트워크 모델과 유전알고리즘을 제안하였다.

확률적 사용자 평형을 이루는 모델은 선형제약을 가진 비선형 목적함수를 최소화하는 문제로 정식화하였다. 네트워크 모델에서는 해의 탐색공간을 줄이고 조사된 통행량을 만족시키기 위해서 흐름보존제약을 활용하였다. 목적함수는 흐름보존, 통행발생량, 통행유입량, 조사통행량 등의 제약을 만족하는 링크통행량과, 경로통행배정을 통하여 구한, 확률적 사용자 평형을 이루는 경로통행량을 만족하는 링크통행량의 차이를 최소화하는 것으로 정식화하였다.

제안된 유전알고리즘에서 유전자는 통행분포, 링크통행량, 여행비용계수 등을 나타내는 벡터로 정의하였다. 각 유전자는 목적함수의 값으로 구한 적합도에 따라 평가되며, 병행단체교차와 돌연변이에 의하여 진화한다.

키워드: 확률적 사용자 평형; 통행분포; 통행배정; 유전알고리즘

A Genetic Algorithm for Trip Distribution and Traffic Assignment in a Stochastic User Equilibrium

ABSTRACT

A network model and a Genetic Algorithm(GA) is proposed to solve the simultaneous estimation of the trip distribution and traffic assignment from traffic counts in the congested networks in a logit-based Stochastic User Equilibrium (SUE).

The model is formulated as a problem of minimizing the non-linear objective functions with the linear constraints. In the model, the flow-conservation constraints of the network are utilized to restrict the solution space and to force the link flows meet the traffic counts. The objective of the model is to minimize the discrepancies between the link flows satisfying the constraints of flow-conservation, trip production from origin, trip attraction to destination and traffic counts at observed links and the link flows estimated through the traffic assignment using the path flow estimator in the logit-based SUE.

In the proposed GA, a chromosome is defined as a vector representing a set of Origin-Destination Matrix (ODM), link flows and travel-cost coefficient. Each chromosome is evaluated from the corresponding discrepancy, and the population of the chromosome is evolved by the concurrent simplex crossover and random mutation. To maintain the feasibility of solutions, a bounded vector shipment is applied during the crossover and mutation.

KEYWORDS:

Stochastic User Equilibrium; Trip distribution; Traffic assignment; O-D matrix Genetic Algorithm;

INTRODUCTION

The conventional function of the network observer has been divided into two components: one is the estimation of an Origin-Destination Matrix (ODM) from traffic counts and the other is the assignment of the ODM to the network to estimate the unmeasured link flows, the location of congestion, the path flows or the path travel times. The problem with this divided approach is that fixed assignment proportions are generally assumed during ODM estimation, which may not be consistent with the congestion-dependent assignment [1].

In un-congested networks, existing models assume that the route-choice proportions are given or predetermined by a congestion-independent, proportional traffic assignment. The entropy maximization or information minimization approaches [2, 3], maximum likelihood methods [4-7] and generalized least squares methods [8-12] have been used in this kind of model.

In congested networks, the estimations of the travel-cost coefficient and the ODM affect each other through travelers' route choices and congestion effects, so the ODM estimation and the traffic assignment cannot be separated. This has led to the use of bi-level programming, whereby the ODM estimation and the traffic assignment sub-problems are solved in sequence [13-15]. A more satisfactory approach was the linear programming Path Flow Estimator proposed by Sherali et al. [16] to estimate user equilibrium path flows, which may then be aggregated to user equilibrium link flows.

Liu et al. [17] made an attempt to estimate the ODM and the travel-cost coefficient θ of the logit-based route-choice probabilities from link traffic counts on un-congested networks. A two-stage heuristic search method was proposed to find the ODM and the coefficient θ simultaneously. They used the observed link flows to calculate the link costs in the logit model, and hence could not solve the inconsistency problem created by the congestion effects of the link flows.

Lo et al. [18] also proposed a procedure for the simultaneous estimation of an ODM and link choice proportions from Origin-Destination(O-D) survey data and traffic counts for congested networks. It is known that the link choice proportions in a network change with traffic conditions and that the dispersion parameter of the route choice model should be updated for a current data set. Their procedure performs ODM estimation and traffic assignment alternately until convergence, in order to obtain the best estimators for both the ODM and the link choice proportions, which are consistent with the survey data and traffic counts.

Recently, Yang et al. [19] developed an optimization model for the simultaneous estimation of an ODM and a travel-cost coefficient for congested networks in a logit-based Stochastic User Equilibrium (SUE). Their model is formulated in the form of a standard differentiable, nonlinear optimization problem with analytical stochastic user equilibrium constraints. They derived explicit expressions of the derivatives of the logit-based SUE constraints with respect to origin-destination demand, link flow, and travel-cost coefficients. The derivatives are computed through a stochastic network-loading approach and then applied to a successive quadratic-programming algorithm to estimate the simultaneous model. They suggested that some further works such as a genetic algorithm approach or a simulated annealing approach are needed to overcome the non-convex property of the problem.

In this paper, a Genetic Algorithm (GA) is proposed to solve the simultaneous estimation of the trip distribution and traffic assignment in the congested networks in a logit-based SUE. The model is formulated as a minimization problem of the linear constrained non-linear objective function. In the model, the flow-conservation constraints of the network are applied to restrict the solution space and to force the link flows meet the traffic counts at observed links. The objective of the model is to minimize the discrepancies between the link flows satisfying the constraints of flow-conservation, trip production from origin, trip attraction to destination and traffic counts at observed links and the link flows estimated through the traffic assignment using the path flow estimator in the logit-based SUE.

In the proposed GA, a chromosome is defined as a vector in real space to represent a set of trip distribution (ODM), link flows and travel-cost coefficient. Each chromosome is evaluated from the corresponding discrepancies and the population of the chromosome is evolved through the concurrent simplex crossover operation and random mutation.

In the next section, some definitions and formulations of the model are introduced. In Section 2, the conventional procedure and techniques of GA are introduced. In Section 3, the procedure of solution method using GA to solve the simultaneous estimation problem in logit-based SUE is presented. A numerical example is provided in Section 4 and general conclusions are summarized in Section 5

1. MODEL FORMULATION

1.1 The Logit-based SUE

Let $G=(N,A)$ be a directed transportation network that consists of a set N of nodes and a set A of directed links. Each link $a \in A$ has an associated flow-dependent travel-cost function $t_a(v_a)$ representing the cost per unit flow or average cost on each link. The $t_a(v_a)$ function can be any positive function of link flow v_a, $a \in A$. Let W denote the set of O-D pairs, q_w denote the demand of O-D pair $w \in W$, R_w denote the set of all routes between O-D pair $w \in W$, and R denote the set of all the routes in the network, $R = \bigcup_{w \in W} R_w$. Let $\mathbf{v} = (\ldots, v_a, \ldots)$ and $\mathbf{q} = (\ldots, q_w, \ldots)$, denote vectors of link flows and ODM respectively. Then the distribution of logit-based SUE trips over the network is described by the following multinomial logit model:

$$P_r^w(\theta, \mathbf{v}) = \frac{\exp\{-\theta(\sum_{a \in A} \delta_{ar}^w t_a(v_a))\}}{\sum_{k \in R_w} \exp\{-\theta(\sum_{a \in A} \delta_{ak}^w t_a(v_a))\}}, \theta \geq 0, r \in R_w, w \in W \qquad (1)$$

where θ is the travel-cost coefficient or route-choice dispersion coefficient in the logit-based SUE model, and δ_{ar}^w is the path-link incidence indicator. That is, if route r between O-D pair w uses link a, $\delta_{ar}^w = 1$, otherwise, $\delta_{ar}^w = 0$, and $P_w^r(\theta, \mathbf{v})$ is the path-choice probability that a user traveling between O-D pair w takes route $r \in R_w$. The coefficient $\theta(\theta \geq 0)$ measures the sensitivity of the route-choice to travel cost. As θ approaches zero, the probability of choosing each route become equal and as θ increases to infinity, the route-choices become extremely concentrated on the least cost route. Thus, the value of the coefficient θ influences the prediction of link flows and the estimation of ODM from link traffic counts.

According to the definition of SUE [20], the logit-based SUE problem is to find a set of path flows and link flows that satisfy the following equations for a given ODM $\mathbf{q}$ and travel-cost coefficient θ.

$$f_r^w = P_r^w(\theta, \mathbf{v}) q_w, \quad r \in R_w, w \in W, \qquad (2)$$

$$v_a = \sum_{w \in W} \sum_{r \in R_w} f_r^w \delta_{ar}^w, \quad a \in A \qquad (3)$$

The logit-based SUE conditions are characterized by combining Equations (2) and (3) in terms of link flows. Namely, the logit-based SUE problem seeks to determine a link flow vector $\mathbf{v}$, that solves

$$v_a - \sum_{w \in W} \sum_{r \in R_w} q_w P_r^w(\theta, \mathbf{v}) \delta_{ar}^w = 0, \quad a \in A \qquad (4)$$

Equations (1)~(3) mean that, when a set of link flows $\{v_a, a \in A\}$ and a travel-cost coefficient θ are given, a set of path-choice probabilities $\{P_w^r(\theta, \mathbf{v}), r \in R_w, w \in W\}$ is determined according to Equation (1), and the path flows are assigned by the set of path-choice probabilities using the Equation (2), and the link flows are collected from the path flows using the Equation (3). Hereby, Equation (4) means that if the set of collected link flows are same as the set of given link flows $\{v_a, a \in A\}$, then the two sets of link flows are the logit-based SUE. While the path-choice probabilities are calculated with respect to a travel-cost coefficient θ as well, so the link flow pattern should be a logit-based SUE with respect to a travel-cost coefficient θ.

1.2 Network Flow Constraints

The model is considered for a static network and all observations and determinations are long-term based. It may be assumed that the link flow $\mathbf{v}$ and the O-D demand $\mathbf{q}$ satisfy the network flow-conservation constraints. Under this assumption, some modifications to the original road network are made for the convenient manipulation of the network flow-conservation constraints.

Let h_a and l_a denote the head and tail node of link $a \in A$, and o_w and d_w denote the origin and destination node of O-D pair $w \in W$, and W_o and W_d denote the set of origin and destination nodes of the O-D pair, respectively.

In the original network $G = (N, A)$, we add an artificial link a for each O-D pair $w \in W$, such that $h_a = o_w$ and $l_a = d_w$. Let A_W denote the set of the artificial links, and then the original network is modified to $G = (N, A \bigcup A_W)$. Then, while the network flow-conservation is satisfied, the O-D demand q_w is the same as the flow of the artificial link a corresponding to the O-D pair $w \in W$.

Let $\bar{o}_k$ and $\bar{d}_k$ denote the amount of traffic produced from node k and the amount of traffic attracted to node k, respectively. The amount of $\bar{o}_k$ and $\bar{d}_k$ can be determined by traffic counts or through the trip generation procedures.

Let $\bar{A} \subseteq A$ denote a subset of links those are observed for the traffic counts, and $\bar{v}_a$ be the traffic count of the observed link $a \in \bar{A}$. Let $\bar{W} \subseteq W$ denote the subset of O-D pairs of which the demands are already known or targeted through network survey or any other prior information, and let $\bar{q}_w$ be the historically known or targeted demand of O-D pair $w \in \bar{W}$. The set $\bar{W}$ is empty if there is no O-D pair of which the demand is known or targeted.

Then we have network flow constraints as follows

$$\left(\sum_{h_a=k} v_a + \sum_{o_w=k} q_w\right) - \left(\sum_{l_a=k} v_a + \sum_{d_w=k} q_w\right) = 0, \quad k \in N \qquad (5)$$

$$\sum_{o_w=k} q_w = \bar{o}_k, \quad k \in W_o \qquad (6)$$

$$\sum_{d_w=k} q_w = \bar{d}_k, \quad k \in W_d \qquad (7)$$

$$v_a = \bar{v}_a, \quad a \in \bar{A} \qquad (8)$$

$$q_w = \bar{q}_w, \quad w \in \overline{W} \qquad (9)$$

$$v_a \geq 0, \quad a \in A \qquad (10)$$

$$q_w \geq 0, \quad w \in W \qquad (11)$$

Equation (5) is flow-conservation constraint at each node. Equation (6) restricts the amount of traffic produced from an origin, and Equation (7) restricts the amount of traffic attracted to a destination. Equation (8) restricts the feasible link flow on the observed links to the observed traffic counts, and Equation (9) restricts the O-D demand to the target O-D demand when it is available. Equations (10) and (11) are the non-negativity conditions. Let Θ denote the solution set of Equations (5)~(11). Then a feasible solution of Θ is an ODM and link flows satisfying the trip production and attraction amount, traffic counts and flow-conservation constraints. Therefore, if a feasible solution of Θ satisfy the condition of the logit-based SUE as well, then it simultaneously solves the trip distribution and traffic assignment those are meeting traffic counts in logit-based SUE.

The traffic counts observed on the network and the target O-D demands might be inaccurate and inconsistent with each other for several reasons, such as counting errors, time discrepancies, dilute identification of zones, etc. When the traffic counts and the target O-D demands are accurate and consistent to the flow-conservation constraints, Θ will have the feasible solutions. If they are inaccurate or inconsistent with each other, then Θ may have no feasible solution. At current stage in this paper, Θ is assumed non-empty.

1.3 The Simultaneous Estimation Model

Let $(\mathbf{q}, \mathbf{v})$ be a given set of ODM and link flows those are feasible in the solution set Θ, and let $\theta \geq 0$ be a given travel-cost coefficient. If the feasible ODM and link flows $(\mathbf{q}, \mathbf{v})$ are consistent with the logit-based SUE, Equation (4) will be satisfied where $P_w^r(\theta, \mathbf{v})$ is the path-choice probability derived by Equation (1). If Equation (4) is not satisfied, the feasible ODM and link flows $(\mathbf{q}, \mathbf{v})$ are not consistent to the logit-based SUE. Let $\hat{\mathbf{v}} = (\ldots, \hat{v}_a, \ldots)$ denote the link flows estimated from the given feasible ODM and link flows $(\mathbf{q}, \mathbf{v})$ by Equation (1)-(3). Then the model for the simultaneous estimation is defined as follows

$$\min F(\mathbf{q}, \mathbf{v}, \theta) = F_1(\mathbf{v}, \hat{\mathbf{v}}), \qquad (12)$$

$$\text{where } \hat{v}_a = \sum_{w \in W} \sum_{r \in R_w} q_w P_r^w(\theta, \mathbf{v}) \delta_{ar}^w, \quad a \in A, \quad (\mathbf{q}, \mathbf{v}) \in \Theta \qquad (13)$$

$F_1(\mathbf{v}, \hat{\mathbf{v}})$ implies the function of discrepancies between the feasible link flows $\mathbf{v}$ and the link flow estimates $\hat{\mathbf{v}}$ derived from $(\mathbf{q}, \mathbf{v}, \theta)$ by Equation (13). If the condition of the logit-based SUE is satisfied, that is, Equation (4) is satisfied, the value of $F_1(\mathbf{v}, \hat{\mathbf{v}})$ become zero. When Θ has no feasible solution, we can relax all or some of the Equations (8) and (9) so that Θ has a feasible solution. Let $\overline{A}_R$ and $\overline{W}_R$ be the set of links corresponding to the relaxed equations, and redefine Equations (8)' and (9)' as follows

$$v_a = \overline{v}_a, \quad a \in \overline{A} / \overline{A}_R \qquad (8)'$$

$$q_w = \overline{q}_w, \quad w \in \overline{W} / \overline{W}_R \qquad (9)'$$

Let Θ' denote the solution space of Equations (5)~(7),(8)',(9)',(10) and (11). Then an objective function for the simultaneous estimation is redefined as follows

$$\min F(\mathbf{q}, \mathbf{v}, \theta) = F_1(\mathbf{v}, \hat{\mathbf{v}}) + F_2(\mathbf{v}, \overline{\mathbf{v}}) + F_3(\mathbf{q}, \overline{\mathbf{q}}), \qquad (14)$$

$$\text{where } \hat{v}_a = \sum_{w \in W} \sum_{r \in R_w} q_w P_r^w(\theta, \mathbf{v}) \delta_{ar}^w, \quad a \in A, \quad (\mathbf{q}, \mathbf{v}) \in \Theta' \qquad (15)$$

The second function $F_2(\mathbf{v}, \overline{\mathbf{v}})$ implies the discrepancies between the feasible link flows of $(\mathbf{q}, \mathbf{v})$ and the link traffic counts corresponding to relaxed constraints. The third function $F_3(\mathbf{q}, \overline{\mathbf{q}})$ implies the discrepancies between the feasible ODM of $(\mathbf{q}, \mathbf{v})$ and the target O-D demands corresponding to the relaxed constraints.

In the GA procedure, the functions can be of any form so that they can appropriately measure the degree of discrepancy between the trip distribution and the trip assignment of the logit-based SUE. The functions do not need to be convex or differentiable which is essential for the solution procedures. In this example, the following weighted Euclidean distance functions are used.

$$F_1(\mathbf{v},\hat{\mathbf{v}}) = \sum_{a \in A} (v_a - \hat{v}_a)^2 \quad (16)$$

$$F_2(\mathbf{v},\bar{\mathbf{v}}) = \sum_{a \in \bar{A}_R} (v_a - \bar{v}_a)^2 \quad (17)$$

$$F_3(\mathbf{q},\bar{\mathbf{q}}) = \sum_{w \in \bar{W}_R} (q_w - \bar{q}_w)^2 \quad (18)$$

These functions correspond to the model of generalized least squares estimation under the SUE constraint for congested networks [11].

2. THE GENETIC ALGORITHM

2.1 Basic Genetic Algorithm

GA is a powerful technique developed by John Holland (1975) over the course of 1960 s and 1970 s. David Goldberg provided significant contributions that increased the popularity of this algorithm, since he was able to solve a difficult problem involving the control of gas-pipeline transmission as his dissertation [21].

As Goldberg stated, GA is different from normal optimization and search procedure. GA is the stochastic search that mimics the process of natural selection. It does not require derivative information and does not require continuity, differentiability, uni-modal, convexity, etc., which may not be satisfied by many real-world problems. It simultaneously searches from multiple starting points and is able to locate the optimal solution.

2.2. Simplex Crossover Operator in a Real-coded GA

In a real-coded GA, each chromosome is coded as a vector representing a point in n-dimensional space. In a GA using the chromosomes coded as binary or decimal strings, crossover can be a single point crossover, a multi-point crossover or a uniform crossover. These kinds of crossover are not applicable, and even if it is applicable, it may lead to slow convergence and poor performance in a real-coded GA. For the real-coded GA, the simplex crossover operator was introduced by Renders et al [22]. The simplex method has been known as a local search technique that uses the evaluation of the current dataset to determine the promising search direction, which was first introduced by Spendley et al. [23].

A simplex is defined by a number of $n+1$ points in an n-dimensional space. The simplex method searches for an optimal point by evaluating a set of points forming a simplex, and continually forming new simplexes by replacing the worst point in the simplex over the centroid of remaining points. The simplex crossover operation begins by choosing $n+1$ chromosomes, representing the points $p_1, p_2, ..., p_{n+1}$ in n-dimensional space. Let p_{n+1} be the point with the worst evaluation. Simplex crossover computes the centroid of the points except the worst point p_{n+1}, $p^c = \frac{1}{n}(p_1 + p_2 +, ..., +p_n)$. The worst p_{n+1} is then reflected across p^c to obtain a new point $p'_{n+1} = p^c + (p^c - p'_{n+1})$. Then the worst point p_{n+1} is replaced by the new point p'_{n+1}. The vector of corresponding chromosomes is changed correspondingly.

2.3 Concurrent Simplex Method

A variant of the regular simplex method is a concurrent version [24] which is similar to the sequential simplex. The variant begins with $n+m$ points, instead of $n+1$ points, in the simplex, where $m > 1$ and is generally significantly greater than one. As the basic simplex method, the best n points $p_1, p_2, ..., p_n$ are selected and their centroid p^c is calculated. Instead of reflecting only one point p_j across p^c, multiple points $p_{n+1}, p_{n+2}, ..., p_{n+m}$ are reflected across p^c to produce $p'_{n+1}, p'_{n+2}, ..., p'_{n+m}$. All new points are re-evaluated, a new set of best points $p_1, p_2, ..., p_n$ are selected, and the process is repeated. The benefit of the concurrent simplex is that it can explore multiple search frontiers simultaneously.

The concurrent simplex crossover may be implemented in the GA. Let P be the population size, N be the number of elitist and S be a number between N and P. Then, N chromosomes in the present generation are allowed to survive into the next generation directly. The best S chromosomes are selected to produce new $S-N$ chromosomes by the concurrent simplex method. The last $P-S$ chromosomes are produced using conventional crossover procedures of randomly chosen parents from P chromosomes. Then the chromosomes of the next generation are completed.

3. GENETIC ALGORITHM FOR THE SIMULTANEOUS ESTIMATION MODEL

In the context of simultaneous estimation of trip distribution and traffic assignment, GA can be applied as follows:

Step 0. Define the parameters of GA (population size, crossover and mutation ratio, etc.)

Step 1. Generate sets of vectors of ODM and link flows of which the value of each element falls within the feasible range of Θ. The number of sets of vectors is equal to the population size.

Step 2. Evaluate the fitness of each set of vectors by estimation the link flows in logit-based SUE. Compute the discrepancy measure of (16), (17), (18) between the estimated link flows and the flows in the set of vectors. This value defines the fitness of the solutions.

Step 3. Perform GA operators to generate the new sets of vectors of ODM and link flows.

Step 4. Check the stopping criterion. If it is satisfied, stop; otherwise go to step 2.

The overall procedure of the proposed GA consists of the operations of initialization, evaluation, crossover, reproduction and mutation.

3.1. Initialization

Obtain any feasible interior point solutions $(\mathbf{q}_i, \mathbf{v}_i), i = 1, \ldots, P$ of Θ (or Θ' in the infeasible case). A two-phase primal affine scaling algorithm is used to obtain an initial interior point solution. Then simplex transformation and random vector projections are used to obtain the feasible solutions scattered randomly in Θ. Let $\bar{\theta}$ be the maximum expected value of θ. Obtain the random scalar values of $0 \le \theta_i \le \bar{\theta}, i = 1, \ldots, P$. Then, the initial populations of chromosomes are the vectors of $p_i = (\mathbf{q}_i, \mathbf{v}_i, \theta_i), i = 1, \ldots, P$.

3.2 Evaluation

Each chromosome $p_i = (\mathbf{q}_i, \mathbf{v}_i, \theta_i), i = 1, \ldots, P$ is evaluated using the objective function (12), (or (14) in the infeasible case). At first $P_w^r(\theta_i, \mathbf{v}_i)$ is calculated from Equation (1) and then $F(\mathbf{q}_i, \mathbf{v}_i, \theta_i)$ is calculated by the objective function (12), (or (14) in the infeasible case). The fitness value of chromosome p_i is set to $1/F(\mathbf{q}_i, \mathbf{v}_i, \theta_i)$. The fitness values are rescaled for the proper selections in the next steps.

3.3 Crossover and Reproduction

Perform the concurrent simplex crossover operation described in Session 2.3 to generate the next generation. The vector of chromosomes resulting from the simplex crossover operation may be an infeasible solution of Θ. To prevent the birth of an infeasible solution, the minimum ratio test and the bounded shift technique are used during the crossover operation.

3.4 Mutation

Perform the random mutation according to the given mutation probability. A projection of random vector to the null space of Θ is added to the vector of the chromosome selected for mutation. The vector of a chromosome resulting from the mutation operation may also be an infeasible solution of Θ. To prevent the birth of an infeasible solution, the minimum ratio test and the bounded shift technique are used during the mutation operation.

3.5 Repeat the Iteration

Repeat the steps of evolution described in Sessions 3.2~3.4 in sequence until certain stop criteria are satisfied. The stop criteria are set in advance of running the procedure. During the iterative procedure, the elitist chromosome is stored and maintained. When the procedure is terminated, the vector of the elitist chromosome remaining in the last generation is the optimal solution.

4. COMPUTATIONAL EXAMPLE

The GA procedure was coded with MATLAB and run on the Windows XP platform. The procedure and the code were tested using a test network provided in the literature [19], which is illustrated in FIGURE 1. The test network consists of 14 links and 9 O-D pairs from origin nodes 1, 2, and 4 to destination nodes 6, 8, and 9, respectively. A set of link traffic counts $\bar{A} = \{6,9,10,11,13\}$ are given. Set $\bar{A}$ constitutes a cut set between the origins and destinations in the network. Though the test network is acyclic and does not have cyclic paths, the proposed GA procedure can solve any general directed networks. The proposed GA procedure does not require an initial seed ODM or a reference ODM, since the initial population is generated randomly.

Equation (16) is used as the objective function for the criteria of discrepancy. TABLE 1 shows travel-cost function, free-flow travel time, capacity, observed flow and the obtained link flow estimates corresponding to the optimal estimated ODM. TABLE 2 shows the given production, attraction from each origin and destination respectively and the optimal estimated ODM obtained through the GA procedure. The amount of trip produced from each origin and the amount of trip attracted to each destination are obtainable through traffic counts or Trip Generation procedures. TABLE 3 shows the list of enumerated paths and the obtained path flow estimates in logit-based SUE corresponding to the optimal estimated ODM. TABLE 4 shows the amount of discrepancy between the obtained link flow and the link flow estimated through trip assignment in logit-based SUE from the optimal estimated ODM.

In the GA, the size of the population and the number of generations to evolve are given as 160 and 50, respectively. FIGURE 2 shows the convergence of the discrepancy value and θ. The optimal value of θ is obtained as $\theta = 1.29034$

5. CONCLUSIONS

A Genetic Algorithm is proposed to solve the simultaneous estimation of the ODM and traffic assignment problem in congested networks using a logit-based SUE. The network flow-conservation constraints are applied to restrict the solution space and to ensure that the estimated trip assignment is consistent with the traffic counts at the observed links. The objective of the model is to minimize the discrepancies between the two sets of link flows. One is the set of link flows consistent with the flow-conservation constraints and the traffic counts. The other is the set of link flows estimated through the traffic assignment in the logit-based SUE. The two sets of link flows are calculated from an ODM and a travel-cost coefficient, those are the subject of simultaneous optimization.

In the proposed GA procedure, each set of an ODM, link flows and a travel-cost coefficient is represented by a real-coded chromosome. The discrepancy value associated with each chromosome is evaluated through the estimation of the link flows. The concurrent simplex crossover operation is used for the proper convergence of the solution to a global optimum. The set of an ODM and a travel-cost coefficient represented by the best chromosome is the solution of the model that optimizes the estimation of the ODM, the traffic assignment, and the travel-cost coefficient simultaneously.

As one of the intrinsic properties of the GA, the proposed procedure does not require the objective functions to be convex and differentiable. The procedure can also solve trip distribution models such as entropy maximization or information minimization, maximum likelihood and generalized least squares only with minor modifications.

REFERENCES

1. Bell, M.G.H., et al., *A stochastic user equilibrium path flow estimator.* Transportation Research Part C: Emerging Technology, 1997. **5**(3-4): p. 197.
2. Van Zuylen, H.J. and L.G. Willumsen, *The most likely trip matrix estimated from traffic counts.* Transportation Research Part B-Methodological, 1980. **14B**(3): p. 281-93.
3. Willumsen, L.G. *Estimating time-dependent trip matrices from traffic counts.* in *9th International Symposium on Transportation and Traffic Theory.* 1984. Delft University,The Netherlands.

4. Spiess, H., *A maximum likelihood model for estimating origin-destination matrices.* Transportation Research Part B: Methodological, 1987. **21**(5): p. 395.
5. Maher, M.J., *Inferences on trip matrices from observations on link volumes: a Bayesian statistical approach.* Transportation Research Part B-Methodological, 1983. **17B**(6): p. 435-47.
6. Van Aerde, M., H. Rakha, and H. Paramahamsan, *Estimation of Origin-Destination Matrices: Relationship Between Practical and Theoretical Considerations.* Transportation Research Record. n 1831, 2003: p. 122-130 03-2296.
7. Rakha, H., H. Paramahamsan, and M. Van Aerde. *Comparison of Static Maximum Likelihood Origin-Destination Formulations.* in *Proceedings of the 16th International Symposium on Transportation and Traffic Theory (ISTTT16). Transportation and Traffic Theory: Flow, Dynamics and Human Interaction.* 2005.
8. Bellei, G., G. Gentile, and N. Papola, *A within-day dynamic traffic assignment model for urban road networks.* Transportation Research Part B: Methodological, 2005. **39**(1): p. 1.
9. Bell, M.G.H., *The estimation of origin-destination matrices by constrained generalized least squares.* Transportation Research Part B: Methodological, 1991. **25**(1): p. 13.
10. McNeil, S. and C. Hendrickson, *A regression formulation of the matrix estimation problem.* Transportation Science, 1985. **19**(3): p. 278-92.
11. Cascetta, E., *Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator.* Transportation Research Part B: Methodological, 1984. **18**(4-5): p. 289.
12. Nie, Y., H.M. Zhang, and W.W. Recker, *Inferring origin-destination trip matrices with a decoupled GLS path flow estimator.* Transportation Research Part B: Methodological, 2005. **39**(6): p. 497.
13. Yang, H., et al., *Estimation of origin-destination matrices from link traffic counts on congested networks.* Transportation Research Part B: Methodological, 1992. **26**(6): p. 417.
14. Yang, H., *Heuristic algorithms for the bilevel origin-destination matrix estimation problem.* Transportation Research Part B: Methodological, 1995. **29**(4): p. 231.
15. Nie, Y., H.M. Zhang, and D.-H. Lee, *Models and algorithms for the traffic assignment problem with link capacity constraints.* Transportation Research Part B: Methodological, 2004. **38**(4): p. 285.
16. Sherali, H.D., R. Sivanandan, and A.G. Hobeika, *A linear programming approach for synthesizing origin-destination trip tables from link traffic volumes.* Transportation Research Part B: Methodological, 1994. **28**(3): p. 213.
17. Liu, S. and J.D. Fricker, *Estimation of a trip table and the [Theta] parameter in a stochastic network.* Transportation Research Part A: Policy and Practice, 1996. **30**(4): p. 287.
18. Lo, H.-P. and C.-P. Chan, *Simultaneous estimation of an origin-destination matrix and link choice proportions using traffic counts.* Transportation Research Part A: Policy and Practice, 2003. **37**(9): p. 771.
19. Yang, H., Q. Meng, and M.G.H. Bell, *Simultaneous Estimation of the Origin-Destination Matrices and Travel-Cost Coefficient for Congested Networks in a Stochastic User Equilibrium.* Transportation Science, 2001. **35**(2): p. 107-123.
20. Sheffi, Y., *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming*

Methods. 1985, Englewood Cliffs, NJ.: Prentice-Hall.

21. Goldberg, D.E., *Genetic algorithms in search, optimization, and machine learning*. Repr. with corrections. ed. 1989, Reading, Mass.: Addison-Wesley Pub. Co. xiii, 412 p.
22. Renders, J.M. and H. Bersini. *Hybridizing genetic algorithms with hill-climbing methods for global optimization: two possible ways*. 1994.
23. Spendley, W., G.R. Hext, and F.R. Himsworth, *Sequential application of simplex designs in optimization and evolutionary opreation.* Technometrics, 1962. **4**: p. 441-461.
24. Yen, J., et al., *A hybrid approach to modeling metabolic systems using a genetic algorithm and simplex method.* Systems, Man and Cybernetics, Part B, IEEE Transactions on, 1998. **28**(2): p. 173.

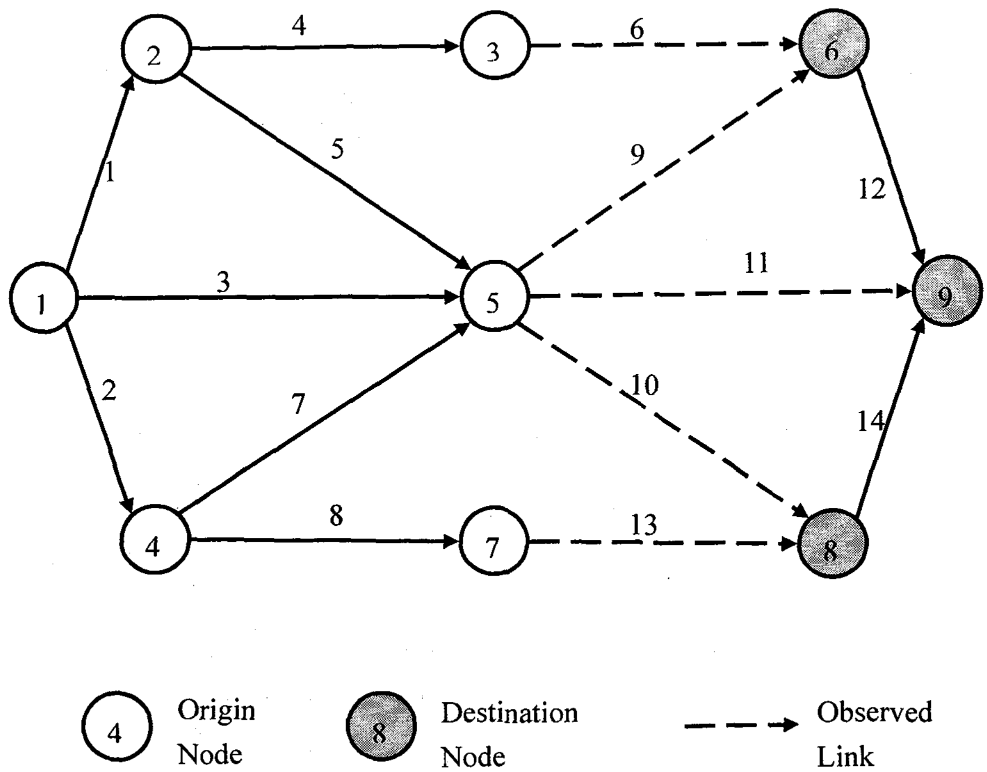

FIGURE 1 The test network for example

TABLE 1 Link performance function, free-flow travel time, capacity, observed link flow and selected link flow

Travel-cost function : $t_a(v_a) = T_a\left\{1.0 + 0.15\left(\frac{v_a}{C_a}\right)^4\right\}$

Link No.	Free-Flow Travel Time(T_a)	Link Capacity (C_a)	Observed Link Flow	Selected Link Flow
1	2.00	280	-	106.0930
2	1.50	290	-	143.5790
3	3.00	280	-	100.3274
4	1.00	280	-	84.9999
5	1.00	600	-	451.0924
6	2.00	300	85	84.9999
7	2.00	500	-	243.5788
8	1.00	400	-	284.9995
9	1.50	500	300	299.9995
10	1.00	700	360	359.9994
11	2.00	250	135	134.9998
12	1.00	300	-	54.9999
13	1.00	350	285	284.9995
14	1.00	220	-	174.9997

TABLE 2 Given production, attraction and the selected optimal ODM estimate

	Selected ODM estimate			
Node	6	8	9	Production
1	67.289	66.511	216.2	350
2	161.23	203.22	65.545	430
4	101.48	200.26	83.255	385
Attraction	330	470	365	1165

TABLE 3 List of paths and estimated path flow in logit-based SUE

Origin	Destination	Path(Link list)	Estimated Path Flows
1	6	1-4-6	12.2102
		1-5-9	21.1601
		3-9	22.4741
		2-7-9	11.4446
1	8	1-5-10	19.5634
		1-3-10	1.5609
		2-7-10	10.5810
		2-8-13	34.8053
1	9	1-4-6-12	8.7365
		1-5-9-12	15.1402
		1-5-11	29.0046
		3-9-12	16.0804
		3-11	30.8056
		3-10-14	29.0664
		2-7-9-12	8.1887
		2-7-11	15.6873
		2-7-10-14	14.8016
		2-7-8-13-14	48.6885
2	6	4-6	58.9943
		5-9	102.2361
2	8	5-10	203.2239
2	9	4-6-12	7.1358
		5-9-12	12.3662
		5-11	23.6902
		5-10-14	22.3527
4	6	7-9	101.4800
4	8	7-10	46.6882
		8-13	153.5765
4	9	7-11	16.4951
		7-10-14	15.5638
		8-13-14	51.1957

TABLE 4 The selected link flow, estimated link flow and discrepancy value

Link No.	Selected Link Flow	Estimated Link Flow.	Discrepancy	Squared Discrepancy
1	106.0930	107.3759	-1.29	1.6641
2	143.5790	144.1970	-0.62	0.3844
3	100.3274	99.9874	0.343	0.117649
4	84.9999	87.0768	-2.077	4.313929
5	451.0924	448.7375	2.35	5.5225
6*	84.9999	87.0768	-2.077	4.313929
7	243.5788	240.9304	2.65	7.0225
8	284.9995	288.2660	-3.27	10.6929
9*	299.9995	310.5703	-10.57	111.7249
10*	359.9994	363.4020	-3.4	11.56
11*	134.9998	115.6829	19.32	373.2624
12	54.9999	67.6477	-12.648	159.9719
13*	284.9995	288.2660	-3.27	10.6929
14	174.9997	181.6688	-6.67	44.4889
Sum				745.5756

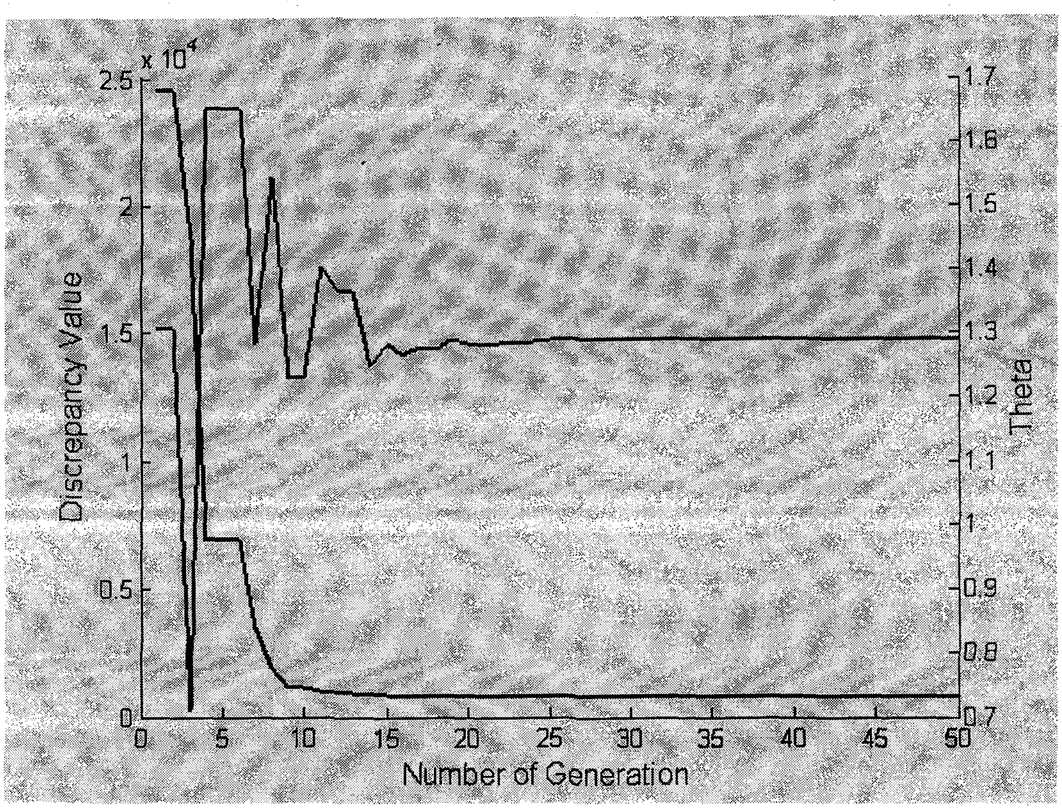

FIGURE 2 Convergence of Solutions